New and Emerging Proteomic Techniques

Dobrin Nedelkov
and Randall W. Nelson
Humana Press, Inc.
Totowa, NJ, USA 07512
ISBN 1-58829-519-2, ISBN 1-59745-026-X
Series on Methods in Molecular Biology,
Volume no. 328, 2006
Hardcover, $99.50 US, 232 pp.

Reviewed by Constance J. Jeffery
Department of Biological Sciences,
University of Illinois at Chicago,
Chicago, IL 60607, USA.
E-mail: cjeffery@uic.edu

New and Emerging Proteomics Techniques contains detailed descriptions of modern proteomics techniques written by experts in the field. The fifteen chapters describe improved versions or novel combinations of proteomics methods, including microarrays, two-dimensional electrophoresis, mass spectrometry, surface plasmon resonance, and bioinformatics, which enable high throughput, increased sensitivity, or other advantages over previous versions or an individual method. Each chapter begins with several pages of background information and/or an overview of the method and its uses. The core of each chapter is a detailed, step-by-step procedure for performing each technique. Most chapters also include an example of an experiment using that technique with sample results. Each chapter concludes with additional detailed notes, including tips for troubleshooting and ways to avoid problems, followed by references. A main benefit of the book is that the procedures and notes are far more detailed than would be found in most journal articles.

The first few chapters describe novel methods of microarrays, which enable assaying hundreds or thousands of proteins simultaneously for high-throughput studies of protein–protein interactions, drug binding, or enzyme activity or for identifying proteins within a mixture.

Chapter one details a method to address perhaps the largest problem in protein microarrays, how to produce arrays of thousands of proteins without having to express and purify each protein. With the nucleic acid programmable protein array (NAPPA) method, the proteins are expressed through in vitro transcription and translation in situ from plasmids immobilized on an array. A plasmid encoded GST tag at the protein's C-terminus is used to attach the newly synthesized protein onto the array via anti-GST antibodies also immobilized on the array. The method enables preparation of the many proteins just in time for use, helping to decrease problems of protein stability. The chapter begins by describing issues in constructing protein arrays and then provides details in preparation of the DNA, arraying the plasmid and anti-GST antibodies, and expression of the proteins in situ. Figures provide an outline of the method and a picture of the proteins expressed on the array.

Micro-arrays of antibodies provide a powerful method for detecting many specific proteins in small amounts of biological fluids or cell extracts, for example for diagnostic purposes. Chapters 2 and 3 describe methods for high sensitivity detection of protein antigens bound to antibody arrays. Chapter 2 details use of rolling circle amplification (RCA). In the example given, antibodies are first bound to the array, they are used to retrieve specific antigens from a mixture of proteins, and a second antibody is used to label the bound antigen. A DNA primer covalently bound to the second antibody primes rolling circle amplification of a DNA circle, and the synthesized DNA is detected by binding of a fluorescent oligonucleotide. Directions and numerous helpful tips are given for the preparation of the DNA circles and antibody-primer covalent adducts, printing of the antibody microarrays, the immunoassay steps, and the imaging and data analysis.

Chapter 3 describes resonance light-scattering particles for the staining of biotinylated detection antibodies. Resonance light-scattering particles are colloidal gold particles coated with an anti-biotin antibody and can be observed at high sensitivity by light scattering. Directions for preparation of the arrays, analysis of antigen samples, imaging and data analysis, and development and optimization of a novel array are given in detail. Figures of the software for interpreting the array are provided.

Chapter 4 details a method for using chemistry-based functional proteomics for profiling active enzymes, with cellular proteases as an example. The chapter begins with a good background on the importance of proteases with examples of cellular roles, variety of activities, and mechanisms of regulation, which provides a great example of the need for a method like this to determine which proteins are active and not just the relative levels of protein expression in different cell types. The method involves the use of tagged selective and irreversible inhibitors of the specific proteolytic activities chosen for study. The protocols cover both the chemical synthesis and the application of several of these inhibitors in the study of active proteases in cell lysates. Labeling of the proteases is followed by separation of the proteins by 1-D and 2-D electrophoresis and Western blotting or autoradiography.

Chapter 5 gives a thorough discussion of the difference 2-D polyacrylamide gel-based approach to protein

Published online October 4, 2006

1044-0305/07/$32.00
doi:10.1016/j.jasms.2006.08.017
(J Am Soc Mass Spectrom 2007, 18, 162–164)

profiling. This is an updated version of one of the original proteomics methods, 2-D gels. The difference method uses a different protein-labeling dye for each protein mixture being assayed. Then the proteins that differ between the two samples can be identified and quantified by analyzing them on a single gel (or multiple gels using the same internal standard), allowing perfect overlay of the resulting images. The chapter provides a great introduction including benefits, range, and limitations of the method, followed by a clear and detailed protocol. The chapter concludes with extensive, useful notes for sample preparation (including two tables listing effects of various reagents on dye labeling), labeling of samples with dye, removal of contaminants such as nucleic acids, range, and limitations. Also included are some considerations to maximize success for some specific applications, such as samples that need detergent or reduction of a sample's complexity for identification of less abundant proteins. The author includes one figure of an example of a scan of a gel with a Cy3 labeled sample mixed with a Cy5 labeled sample and also references several papers with additional examples.

Chapter 6 describes a method of determining the molecular mass and oligomeric state of proteins by immediately following fractionation of the sample by size-exclusion chromatography with analysis using three detectors in series for measurement of static laser light scattering (LS), refractive index (RI), and ultraviolet absorbance (UV). The combined method (SEC-UV/LS/RI analysis) can be used for native soluble proteins as well as challenging samples such as glycosylated proteins, proteins with other posttranslational modifications, or membrane proteins solubilized in nonionic detergents (protein detergent complexes), for which oligomeric state determination is not always possible by SEC alone. In addition, if the chemical nature of a carbohydrate, lipid, detergent, or other nonpeptide moiety is known, the combined approach can be used to estimate its amount. The introduction provides an explanation of the calculations involved with examples of results with two well characterized membrane proteins of known structures, a heptameric hemolysin and a trimeric maltoporin. The methods, notes, and a table detail setup and calibration of the instruments, sample preparation and measurements, selection of the optimal SEC column, choice of protein standards, and values for refractive index of the polypeptide chain, and potential bound moieties. Also included is a step-by-step description of how to use the results for calculating the molar mass of the polypeptide, oligomeric state, amount of the nonpolypeptide moiety, and molar mass for the whole complex.

Chapters 7 and 8 describe methods that involve surface plasmon resonance (SPR) imaging, which can be used to study the binding of proteins to other molecules without the need to label the proteins. Chapter 7 begins with an explanation of how SPR works followed by detailed methods for preparation of oligonucleotide and peptide arrays for measuring DNA–protein interactions and peptide–protein interactions, respectively. Figures illustrating the DNA array fabrication process and assembly of the SPR sample holder/flow cell are helpful. Description of the imaging methods is accompanied by figures of the results for two DNA binding proteins (response regulators VanR and OmpR) binding to an array of cognate and control oligonucleotides and results for S protein binding to and array of native and S peptide variants. In Chapter 8, surface plasmon resonance imaging is combined with mass spectrometry (SPR-MS), also referred to as biomolecular interaction analysis mass spectrometry. By combining the two methods, the same proteins that are retrieved by affinity to antibodies, oligonucleotides, or peptides on the array are immediately analyzed further by mass spectrometry. The combination enables (1) the identification of structural features of the interacting proteins, such as mutations or posttranslational modification via measurement of their masses, (2) assessment of quantitative variations in protein interactions, and (3) kinetics of protein interactions. The combination of methods is illustrated by an example using immobilized antibodies to capture degraded cardiac troponin I fragments from a mixture containing human serum proteins. Notes include helpful information, such as an explanation of the need for higher chip ligand density in this method than in usual SPR.

As in Chapter 8, Chapter 9 combines mass spectrometry analysis with a method to enrich a specific protein target from a mixture. The affinity mass spectrometry (AMS) described is a high-throughput method that makes use of microaffinity pipet columns in a parallel processing robotics workstation. The methods are illustrated with isolation and characterization of cystatin *C* from human plasma. Methods described include derivatation of affinity pipets (columns) with antibody, affinity purification of the protein of interest, and MS analysis. Figures help by illustrating the setup of the columns and MALDI-target and illustrate the results for 96 human samples, including representative individual mass spectra for a sample from an average individual and for samples from individuals with additional truncation variants or a point mutation.

Chapter 10 adds isotope-coded affinity tags (I-CAT) to affinity chromatography and mass spectrometry for identifying relative amounts of different proteins in biological samples. For example, it can be used to look for changes in the amount of different proteins in healthy and diseased tissues. The method involves tagging of cysteine residues in the proteins in the two biological samples to be compared with two versions of a tag containing different stable isotopes (that differ by 9 Da). After tagging, the two protein mixtures are combined and digested with trypsin. The resulting peptides are subjected to affinity chromatography to collect the tagged cysteine containing peptides, which are then analyzed by mass spectrometry. The corresponding proteins are identified through database

searches, and their relative level of expression in the two samples is determined by the heavy:light ratio of their peptides. The method described is high throughput, since most proteins contain cysteines, and at the same time the problem of too much complexity in a mixture of proteins is solved by looking only at those peptides that contain cysteine. The description and protocol are clear with the aid of a flow chart of steps. Perhaps a sample trace of the results would have been a nice addition.

Chapter 11 describes multidimensional protein identification technology (MudPIT) for the separation and analysis of complex mixtures of peptides from biofluids, tissues, whole cells, organelles, or protein complexes. It combines biphasic or triphasic microcapillary columns (containing both reversed-phase and strong cation exchange resins) in high-performance liquid chromatography followed by tandem mass spectrometry and database searching. The chapter gives good details about construction of the microcapillary columns, peptide sample preparation, equipment setup, and suggested gradient profiles for the chromatography, with notes on data analysis using SEQUEST or manually. It also includes tips for dealing with clogged columns, avoidance of glycerol, and similar useful notes.

Chapter 12 also describes a method of sample preparation for mass spectrometry. In this case, it is used to isolate peptides that contain N-linked glycosylation sites, which is a common and critical posttranslational modification of many proteins on cell surfaces, secreted proteins, and proteins in biological fluids. As the authors note, many of the specimens that are accessible for diagnostic or therapeutic purposes, including proteins that are biomarkers or drug targets, are glycosylated, and glycosylation is important for the correct folding, localization, and function of many of these proteins. The chapter details attachment of the glycoproteins to solid support via their N-linked glycan, trypsin digestion of the proteins, removal of the nonglycosylated peptides, isotopic labeling of the glycopeptides with stable isotopes (heavy or light succinic anhydride), release of the glycosylated peptides from the solid support by digestion by a glycosidase, and then identification and quantification of the formerly glycosylated peptides by tandem mass spectrometry and database searching.

While Chapter 12 details identification of the peptides with N-linked glycosylation, Chapter 13 describes a novel algorithm, STRucture of OLIGOsaccharides (StrOligo), for using mass spectrometric data to identify the composition and possible structures of the glycan. In the first step, the algorithm identifies the *m/z* value between two fragments differing by only one residue and builds a relationship tree, similar in the method used in protein sequencing by mass spectrometry. Next, the algorithm uses the tree to propose possible compositions of monosaccharides totaling the *m/z* value of the glycans studied, to predict the composition of the glycan. To propose a structure, the algorithm then uses the tree and composition, along with information from biosynthetic pathways, to predict possible structures of the glycan and provide scores as to the match between the structures and the observed data, which reflects the fit with the experimental results. Figures of the interface and samples results are helpful.

Chapter 14 focuses on data analysis for using spectra from matrix-assisted laser desorption/ionization mass spectrometry (MALDI - MS) in biomarker discovery. The chapter details issues in identifying significant peaks that differ between disease and healthy samples, for example collecting of a significant number of spectra, calculating appropriate baselines, and identifying features in the spectra. An example is given using 77 control and 93 disease cases from the National Ovarian Cancer Early Detection Program.

The final chapter describes analysis of tandem mass spectra using the global proteome machine (GPM), an open-source, freely available informatics system for comparing spectra with all of the peptide sequences for a selected organism's proteome. The system can assist in identifying proteins, point mutations, posttranslational modifications, as well as artifact side-chain modifications. It also connects to a repository of experimental results from other laboratories for comparison and genomic information and literature resources. A series of figures and specific directions help guide the user through the software interface. Figures include a list of all results, and the more detailed results for a single protein, which includes a sequence coverage map, the amino acid sequence with indications of exon structure, SNPs, peptide sequences that have been assigned to spectra, expectation value, and a protein validation page showing which peptides have been found by other laboratories for that same protein. The program also provides a spectrum-to-peptide sequence assignment and links to the latest information about a particular protein from a curated genomic or cDNA sequence collection. The directions and notes indicate how the user can customize the calculations by specifying possible side-chain modifications, different cleavage enzymes, point mutations, species, and similar information.

Overall, the book's background information and detailed protocols and notes would be of interest to researchers in many areas who are interested in knowing more about the types of studies now made possible by these powerful, combined or individual proteomics techniques, and how they might be applied to current questions in basic and medical science.